Supporting Information

Buschman et al. 10.1073/pnas.1104666108

SI Materials and Methods

Two male rhesus monkeys, Sp and Si, weighing 13 and 6 kg, respectively, were trained on a visual working memory task (Fig. 1*A*). All procedures followed the guidelines of the Massachusetts Institute of Technology Committee on Animal Care and the National Institutes of Health. Animals were prepared using standard procedures (1, 2). Chambers were stereotaxically placed over frontal and parietal cortices (in the same hemisphere), using structural MRI scans. Unique software was developed in Matlab that produced 3D models of each animal's skull and brain in stereotaxic coordinates. This method allowed accurate placement of electrode penetrations into lateral prefrontal cortex, frontal eye fields, and lateral intraparietal cortex simultaneously. Epoxy-coated tungsten electrodes (FHC) were used for recording as well as for microstimulation. Electrodes were lowered using a custom-built microdrive assembly that lowered electrodes in pairs from a single screw. The microdrive assembly was designed to allow for a high density of electrodes (1-mm spacing) to maximize the number of simultaneously recorded neurons across our three regions of interest. The electrodes were acutely lowered through an intact dura at the beginning of every recording session and allowed to settle for a minimum of 2 h before recording. This process ensured stable isolation of the activity single neurons. After each recording session, the electrodes were retracted and the microdrive assembly was removed from the well.

Spiking activity was recorded across a maximum of 50 electrodes simultaneously. Both spiking activity and local field potentials were referenced to ground. The signal from each electrode was divided into spiking activity and a local field potential by filtering between 154 Hz and 8.8 kHz for spikes and between 3.3 and 88 Hz for the local field potential. Only spiking data were analyzed for this paper. Waveforms were stored and single neural activity was sorted from the raw spiking activity signal off-line using a combination of principal component analysis of waveform traces along with other properties of the recorded waveforms (amplitude, trough/peak latency, etc). An infrared-based eye-tracking system monitored eye position at 240 Hz. Behavioral control of the paradigm was handled by the Monkeylogic program (www.monkeylogic.net) (3, 4). All analysis code was custom written in Matlab or C.

Behavioral Task. The trial was initiated when the animal fixated on a point at the center of the screen. Fixation was required within 1.75° of visual angle (dva) of the fixation point. After a short fixation period (500 ms), the animal was presented with an array of colored squares for 800 ms. A long stimulus period was chosen to ensure the animal had enough time to fully attend to and process all of the items in the array. After the sample period the stimuli were removed and the animal was required to maintain fixation for a memory delay that ranged from 800 to 1000 ms. Following the memory delay, a test array was presented. The test array was identical to the sample array except that the color of a single stimulus was changed. The animal's task was to make a single, direct, saccade to the changed object.

Six new stimulus locations were chosen each day. There were always three locations in each visual hemifield, ranging from ±75 angular degrees from the horizontal meridian and between 4 and 6 dva from fixation. Stimuli were colored squares 1 dva on a side. Two colors were chosen for each location every day. Both the location and colors were changed each day, to prevent monkeys from adopting any long-term memorization strategy. The colors were drawn from a predefined population of 14 colors in a random manner as long as 2 colors were not too difficult to discriminate at a particular location (i.e., red and pink were never paired). This process ensured a large degree of variety in the stimuli used on a particular day and thus required the animals to encode and hold in memory the array presented on each trial and detect its change rather than memorizing fixed stimulus-response associations. The location of the target item (which was the stimulus that flipped from one of the color pair in the sample array to the other in the test array) was chosen randomly on each trial.

The total number of stimuli in the visual array varied from two to five items on each trial. Early in training the number of total stimuli in the array was chosen randomly. However, we noticed both animals showed behavioral evidence of independent capacities in each hemifield (Fig. 2). Anticipating the need to compare trials with one, two, and three objects in a hemisphere, we then equalized the number of trials in each of these conditions. Therefore, the constellation of the stimuli used was pseudorandomly chosen such that both the distribution of trials with one, two, or three items in the target's hemifield and the distribution of total number of stimuli in a trial were flattened. This process did not alter the probability of the target location.

The animals performed a minimum of 720 correct trials during recording sessions, ensuring at least 20 trials for each target, color, and number of ipsilateral stimuli conditions. Both monkeys performed the task well above chance and with similar capacities (Fig. S1). Only trials during which the animal was consistently attempting the task were used: Three of five trials in any given period had to be attempted to be included. Nonattempted trials were failing to initiate the trial or broken fixation before the trial was completed. For the main analysis (Figs. 2*D* and 3), all attempted trials (correct or error) were used, regardless of the target location. For correct/error analyses, only trials where the animal correctly/incorrectly identified a change at a given object were used.

Estimating Information from Behavior. To understand how capacity limitations are reflected in neural activity, we first must estimate the animal's behavioral capacity. An information theoretic approach was used to fully account for chance behavior and make no assumptions about the animal's strategy in solving the task. We were interested in knowing how much information the animal had about a given stimulus display. As the animal uses this information to make its behavioral choice, this is equivalent to asking how much information the animal's choice gives us about where the target is located: $I(\text{target; behavior}) = H(\text{target}) - H(\text{target} \mid \text{choice})$, where $H(x) = -\sum_i^N p(x_i) \times \log_2 p(x_i)$ is the uncertainty of x. The uncertainty of the location of the target, H(target), was determined directly from the likelihood of the target appearing at each possible location in the display (i.e., a flat distribution). The uncertainty of the target, given the animal's choice, was equal to the probability the target was at each position given that the animal chose a particular location.

This process determines the information for a whole stimulus array (Fig. 1*C*), but we were also interested in the information the animal had about stimuli in each hemifield. The animal's behavior and neural activity showed the two hemispheres were independent. Therefore, the information the animal had about the overall display can be decomposed into the sum of the two independent hemifields. This process allows us to estimate the information for the display in each hemifield directly from the total display: $I(\text{display}) = I(\text{left display}) + I(\text{right display})$. For example, given a particular display:

$$I\left(\begin{matrix}X\\X\end{matrix}\middle|X\right)=I\left(\begin{matrix}X\\X\end{matrix}\middle|\right)+\left(X\middle|\begin{matrix}X\\X\end{matrix}\right).$$

As there were 56 unique total stimulus displays, and 8 unique displays for the left and right sides, we can solve the set of linear equations to estimate the information for each hemifield. These estimates are then used to estimate the average information given the number of stimuli in the hemifield (Fig. 2*C* and Fig. S1*B*) and show that both animals had a capacity limit between one and two items in each hemifield. Matlab code for using this information theoretic approach to behavioral analysis is gladly shared upon request.

Determining Whether Behavioral Information Was Capacity Limited. To determine whether the information the animals had about the display (Figs. 1*C* and 2*C*) was capacity limited we use a model selection approach. The observed information was modeled with both a linear curve, $I = \beta_1 x + \beta_0$, and a capacity-limited curve,

$$I=\begin{cases}\beta_1 x+\beta_0, & x<c\\ \beta_1 c+\beta_0, & x\geq c.\end{cases}$$

We compared the goodness-of-fit using both a validation test and the Bayesian information criterion (BIC). For the validation test a randomly selected subset of the trials (70%) was used to estimate the parameters of each model in turn. The predicted values from each model were then compared with the reserved trials to measure the error in the model fit. This process "validates" how well the model captures the data without overfitting. By repeating this process multiple times a distribution of the difference in errors for our two models can be generated. The validated errors for the capacity-limited model were significantly lower than those for the linear model for the information about the total display (Fig. 1*C*, $P = 0.026$). This difference was even greater for the ipsilateral information (Fig. 2*C*, $P = 4.4 \times 10^{-4}$). Alternatively, one can use the BIC to test how well models fit the observed data while correcting for differences in the number of free parameters (5, 6). As with the validation test, the capacity-limited model was found to fit better (have a lower BIC) for both the total ($BIC_{linear} = -24.9$, $BIC_{cap\text{-}lim} = -31.0$) and ipsilateral displays ($BIC_{linear} = -17.4$, $BIC_{cap\text{-}lim} = -21.8$). Both of these tests confirm that the capacity-limited model fits the observed behavioral data better than a simpler linear model, suggesting the monkeys' behavior was indeed capacity limited.

Estimating Capacity from Behavior. Although estimating the information available to the animal from their behavioral performance fully accounts for chance levels and alternative strategies it is difficult to directly compare with previous human psychophysical work. Therefore, we also computed the capacity for each animal using a behavioral model similar to that previously used in human capacity limitation studies. The animals were modeled to correctly remember c objects worth of information on a given trial (their capacity). This limit was not necessarily discrete and partial object information was allowed (e.g., 3.7 items). The animals' baseline performance was taken to be b to compensate for fixation breaks, periods of inattention, etc. Using these two parameters, we modeled their likelihood of getting a trial correct as

$$p(n)=\begin{cases}b, & n\leq c\\ b\dfrac{c}{n}, & n>c,\end{cases}$$

where n is the number of objects in the visual array (either within a hemifield or across the entire visual scene depending on whether we were estimating the hemifield or total capacity). Confidence intervals were estimated by a bootstrapping procedure: a new population of trials (from the >42,000 recorded) was randomly selected (with replacement) and the parameters were reestimated. Repeating this process 1,000 times resulted in a distribution for each parameter from which 95% confidence intervals (95% CI) were determined. However, it should be noted that this model-based approach does not fully accommodate for alternative strategies (such as an exclusionary strategy) or guessing. This problem is typically corrected in human psychophysical work by subtracting the false alarm rate from the hit rate (7), but this process becomes complicated when capacities between hemifields are independent but guessing is not (as the animal can make only one choice). However, our estimate from this simple behavioral model is well aligned with the plateau observed in the information theoretic analysis (which fully compensates for any of these strategies), suggesting the animals were not adopting a complicated behavioral strategy.

Recording Locations. A total of 50 electrodes were implanted into parietal and frontal cortex simultaneously, up to 25 in each anatomical area. Data are presented from 28 sessions (14 each for Si and Sp). A total of 1,334 neurons were recorded across all three anatomical regions in two monkeys (339 from the lateral intraparietal area, LIP; 640 neurons from lateral prefrontal cortex, LPFC; and 355 neurons from the frontal eye fields, FEF). We chose to record from frontal and parietal cortex as both regions have been implicated in working memory (8–11) and capacity limitations (12–14). Furthermore, all three regions have been previously shown to encode stimulus color, particularly when task relevant (15, 16). A total of 540 neurons were recorded from monkey Si and 794 neurons from monkey Sp. Similar behavioral and electrophysiological results were obtained from each animal alone (Figs. S1 and S2), so they are combined for presentation.

The LIP recording well was placed at ~4 mm posterior from the interaural plane and was placed using structural MRIs. To identify LIP neurophysiologically, we trained the animals on a delayed saccade task. During central fixation, a brief spot of light was flashed in the periphery. After a memory delay, the fixation point was extinguished and the animal made a saccade to the remembered location of the light spot. This has been used to isolate LIP from surrounding regions, as it is the only region in the parietal cortex that shows spatially selective memory delay activity (17). The animals performed the delayed saccade task at the beginning of every recording session. Electrodes were considered to be within LIP for that session only if a neuron isolated from that electrode showed memory delay activity selective for the remembered location ($P < 0.05$ using the ωPEV statistic).

The frontal recording well was placed at ~32 mm anterior from the interaural plane. Microstimulation was used to demarcate the frontal eye fields from dorsolateral prefrontal cortex. An Isolated Pulse Stimulator (Model 2100; A-M Systems) was used for electrical stimulation. Stimulation was delivered as a 200-ms train of biphasic pulses with a width 400 μs and an interpulse frequency of 330 Hz, using the same electrodes used for recording. Current level was started at 150 μA and reduced to find the threshold at which an eye movement vector was elicited 50% of the time. Only sites that had thresholds of stimulation amplitudes <50 μA were classified as belonging to the frontal eye fields (18). Anterior sites were classified as belonging to the LPFC. In general, stimulation at LPFC sites did not elicit eye movements even at the highest current amplitude tested (150 μA).

For the analysis of the overall loss of information (Fig. 3), we required each neuron's activity to have been recorded for a minimum of 30 trials for each object location. This restriction yielded a population of 284 LIP neurons, 584 LPFC neurons, and 325 FEF neurons. However, this requirement was relaxed to 15 trials for the correct-alone (Fig. 4*A*) and 5 trials for correct vs. error (Fig. 4*B*) analyses, as these analyses required the analyzed object to be the eventual target. In addition, only selective

neurons (assessed across all trials, see below) were used. This restriction yielded 41 selective LIP neurons, 130 selective LPFC neurons, and 67 selective FEF neurons included in the correct/error analysis (Fig. 4).

Information from Individual Neuron Firing Rates. The factor of interest was the color identity of each object in the array, which was not known to the monkey before the trial began. The key question was how neural information about color identity of the objects changed as the number of objects to be encoded and remembered was increased. We assessed selectivity for the identity of each stimulus for each neuron using a percentage of explained variance (PEV) statistic (for example, neurons, Fig. S2). The PEV reflects how much of the variance in a neuron's firing rate can be explained by the color identity of a particular stimulus. Typically, PEV is expressed as η^2, $\eta^2 = \mathrm{SS}_{\text{Between Groups}}/\mathrm{SS}_{\text{Total}}$, where $\mathrm{SS}_{\text{Total}} = \sum_i^N (x_i - \bar{x})^2$ and $\mathrm{SS}_{\text{Between Groups}} = \sum_{\text{group}}^G n_{\text{group}} (\bar{x}_{\text{group}} - \bar{x})^2$. Unfortunately, for smaller sample sizes, the η^2 statistic has a strong positive bias. Therefore, for all of our statistics in this paper we used the ω^2 statistic (ωPEV),

$$\omega^2 = \frac{\mathrm{SS}_{\text{Between Groups}} - \mathrm{df} \times \mathrm{MSE}}{\mathrm{SS}_{\text{Total}} + \mathrm{MSE}},$$

where df is the degrees of freedom (i.e., the number of groups, G, − 1) and MSE is the mean squared error, $\mathrm{MSE} = \sum_i^N (x_i - \bar{x}_{\text{group}})^2$. ω^2 is an unbiased measure (19), resulting in a zero-mean statistic when there is no information (e.g., see baselines of Fig. S2, *Lower* row). For this reason, and as ω^2 is not computationally expensive to compute, we prefer ωPEV over the more prevalent η^2 form. However, it is important to note that although the mean of ω^2 is unbiased, the distribution of observed values still varies with the number of observations (i.e., the skew of the distribution). Therefore, for all of our comparisons, conditions were balanced for the number of trials in each group. Balancing was accomplished by stratifying the number of trials in each group to a common value: A random subset of trials was drawn (equal to the minimum trial number across groups) and the statistic was calculated. This process was repeated 25 times and the overall statistic was taken to be the mean of the stratified values.

To determine whether and when the observed level of ωPEV was significantly different from chance, we used a randomization test. The association between neural activity and stimulus identity was randomly shuffled and the ωPEV was recalculated. By repeating this process 500 times a null distribution was constructed. The observed ωPEV was then compared with this null distribution to determine the likelihood of the observed ωPEV. The time course of ωPEV was calculated in windows of 100 ms, slid every 10 ms. Neurons were independently tested for selectivity within the sample period (0–800 ms from the onset of the visual array) and the delay period (0–800 ms from the offset of the visual array).

The significance threshold for the amount of information was determined for each window by dividing the typical significance threshold ($P \leq 0.05$) by the number of comparisons made across each window. A neuron was considered selective for object identity if its likelihood reached this significance threshold for two independent windows of time (i.e., two consecutive 100-ms windows). This process corrected for multiple comparisons across time and reduced the chance that the selective responses were due to nonphysiological anomalies.

As used here the ωPEV statistic makes one assumption: Neurons encode stimuli by modulating their average firing rate within the analyzed window of time. Importantly, it does not make any assumption about the consistency of neural response over time or between displays of different sizes. The ωPEV statistic allows us to take an agnostic approach to decoding information about stimulus identity in different conditions (and across time), capturing as much of the selectivity as possible and avoiding any bias across conditions.

Testing for Independent Hemispheres in Neural Activity. As noted in the main text, we found behavioral evidence that the two hemispheres of visual space had independent capacities. To determine whether this effect extended to the neural level, we determined the degree of object selectivity for neurons in LPFC, LIP, and FEF for each display used (as shown in Fig. 2*A*). We then determined the impact of adding objects to the ipsilateral and contralateral hemifields, using a two-way ANOVA. As seen in Fig. 2*D*, LPFC neurons showed a reduction in selectivity with the addition of stimuli in the ipsilateral hemifield but not when adding objects to the contralateral hemifield during the sample period (100–800 ms after sample onset) and memory delay (100–800 ms after sample offset). Similar results were also seen in LIP, LPFC, and FEF for the time periods of interest highlighted in Fig. 4 (150–350 ms for LIP, 300–800 ms in LPFC, 450–800 ms for FEF; P values for LIP and FEF in main text; LPFC, ipsilateral effect, $P = 0.0011$, and contralateral effect, $P = 0.73$).

Testing for Significance on Neuron Population Level. We determined when the amount of average object information across the neuron population reached significance (above baseline) by a nonparametric, paired, Wilcoxon signed rank test. As we examine only neurons whose firing rate carried significant object information, the population will necessarily be biased above zero. To compensate for this bias, we chose the first 50 ms as our baseline period, averaged over all three conditions (i.e., ignoring the number of stimuli in the array). As this time period fell within our time window of selection, any bias due to selecting the neurons alone was corrected. Furthermore, as our baseline is averaged across all three conditions this process should not lead to any bias in our observed timing differences for each condition (see below for this procedure).

Significant differences between below-capacity and above-capacity conditions were determined using a randomization test. Briefly, the association between observed information (in a given 100-ms window) and the number of stimuli on the screen was broken by randomly shuffling conditions. The average difference in population information was determined for each random shuffle. This process was repeated 1,000 times and the observed difference in information was compared with the population of randomized values to determine the likelihood of seeing our observed value by chance. The first time this likelihood fell below 5% for both 1 vs. 2 and 1 vs. 3 was taken as our time-to-first significance. However, as this measure can be biased by statistical power, we used a different statistic to measure the time to significant deviation: the point of maximum rise in the difference function (below capacity minus above capacity). This statistic is explained next.

Details of the Latencies of Neural Information. As noted in the main text, the time point at which selectivity for the identity of an object exceeded baseline depended on whether the displays were above or below capacity. When the display was below the animal's capacity (i.e., one object in the hemifield), object information appeared in LIP at 193 ms after the onset of the visual array (95% confidence interval, 149–229 ms). LIP selectivity was followed by responses in LPFC (317 ms; 95% CI, 249–359 ms) and FEF (291 ms; 95% CI, 249–339 ms). This result is consistent with a bottom–up flow of sensory inputs from posterior to anterior cortex (LIP < LPFC, $P < 10^{-3}$; LIP < FEF, $P = 3 \times 10^{-3}$, randomization test). There was no significant difference between the two frontal regions (FEF < LPFC, $P = 0.23$, randomization test).

Information about the object's identity was seen in the reverse order when the display was above capacity (two or three objects

per hemifield): first in LPFC and FEF, followed by LIP. Significance in LPFC occurred at 315 ms (95% CI, 249–339 ms) and 318 ms (95% CI, 259–359 ms) after visual array onset for two and three items, respectively. FEF was at approximately the same time: 322 ms (95% CI, 289–409 ms) and 290 ms (95% CI, 249–309 ms) for two and three items, respectively. Significant information about two or more objects was not observed in LIP until 527 ms (95% CI, 439–549 ms) and 474 ms (95% CI, 399–539 ms) after the array onset. Information was significantly earlier in LPFC and FEF than in LIP (LPFC < LIP, $P < 10^{-3}$; FEF < LIP, $P < 10^{-3}$, by randomization test for both two and three objects). Again, there was no significant difference between the frontal regions (LPFC < FEF, $P = 0.49$ for two items; FEF < LPFC, $P = 0.23$ for three items).

This pattern of results suggests top–down input may be important for representing object identity information when the visual array is above the animal's capacity. Indeed, the difference in selectivity in LIP was significantly later in above-capacity trials compared with below-capacity trials ($1 < 2$, $P < 10^{-3}$; $1 < 3$, $P < 10^{-3}$, randomization test). However, there was no significant difference in the timing of selectivity for above-capacity conditions in LIP ($3 < 2$, $P = 0.09$) or for any condition in either LPFC ($2 < 1$, $P = 0.38$; $1 < 3$, $P = 0.35$; $2 < 3$, $P = 0.38$, randomization test) or FEF ($1 < 2$, $P = 0.48$; $3 < 1$, $P = 0.30$; $3 < 2$, $P = 0.49$, randomization test).

Loss of Information Between Two- and Three-Item Displays. The shared resource hypothesis predicts that an information source is divided among the objects currently remembered. Therefore, information about a stimulus in a three-item display should be reduced from when that stimulus is in a two-item display. Whereas this pattern is seen late in the delay period for LPFC neurons (Fig. 3*B*, *Right*), no significant difference was observed during the sample period. As noted in the main text, one interesting hypothesis is that this difference in information loss reflects a difference in capacity between these two time periods. For example, when the stimuli are present (or were recently present), the visual trace may provide a source of information that is less limited (or even unlimited). However, this argument extends from a negative finding (that there was no significant difference between two and three objects), making it difficult to interpret. As the expected difference in information per stimulus decreases with added items (i.e., from 1 unit of information compared with 1/2 and then from 1/2 compared with 1/3), our ability to resolve the difference is reduced, possibly leading to an inability to observe a significant difference. However, this concern is mitigated by the fact that we are able to resolve the difference late in the trial.

Loss of Information for Individual Neurons. The loss of information above capacity is clear on a population level (Figs. 3 and 4). However, this loss either could be an effect at the population level (e.g., differential recruitment of neuron groups) or could occur for individual neurons. To investigate this we isolated neurons that were significantly selective ($P < 0.05$) for the same object in both below-capacity and above-capacity displays. Although this process greatly reduced the number of neurons in our population (38 from LPFC, 13 from FEF, and 4 from LIP; Fig. S3), the large majority of neurons in all three regions showed a decrease in selectivity during above-capacity trials in comparison with below-capacity trials (32 of 38 in LPFC, 11 of 13 in FEF, and all 4 in LIP). This difference was significantly above chance across all three areas ($P = 4.0 \times 10^{-8}$) as well as for LPFC ($P = 1.2 \times 10^{-5}$), FEF ($P = 0.011$), and trending for LIP ($P = 0.0625$, limited by small sample size surviving the multiple selection).

Selectivity on Correct and Error Trials. As noted in the main text, we observed a significant reduction in object information in all three areas (for our time periods of interest) when a second object was added to the visual display. To perform this analysis only information about the correctly identified target was used. Unfortunately, this restriction, coupled with the decaying behavior with three items in a hemifield, did not allow for analysis of three-item displays in the correct-only analyses.

In addition to LPFC neurons, FEF neurons also carried significant object information during error trials with two items in the display (average of 0.83% explained variance, bias corrected; $P = 0.024$, permutation test). Although information in error trials trended toward less than during correct trials (which had a 1.2% explained variance, bias corrected) the difference did not reach significance ($P = 0.31$). In contrast, as noted in the main text, LPFC neurons carried significantly less information during error trials than during correct trials ($P = 0.008$). Although this difference may be due to differences in sample size (130 neurons for LPFC, 67 for FEF) it may also reflect a stronger correlation between LPFC neurons and behavior in comparison with FEF, suggesting LPFC plays a larger role in directing behavior during a working memory task. LIP neurons did not carry any significant information about stimuli during two-item conditions (Fig. 4*A*).

As noted in the main text we observed a significant amount of object information in selective LPFC neurons during error trials alone (for both the sample, Fig. 4*B*, and delay, main text). We also tested this across the entire population of LPFC neurons by determining whether there were a significant number of neurons carrying significant object information using error trials alone. As for the correct–error comparison, we limited our analysis to the two-item condition (where there were sufficient error trials). The information criterion was the same used for all trials (i.e., selectivity in two independent 100-ms windows, see above). As before, we found a significant number of LPFC neurons conveying object information during the sample period ($P = 0.0056$) with a trend continuing into the delay period ($P = 0.0534$). This result confirms our original finding of object identity information, even when that object is "forgotten".

1. Buschman TJ, Miller EK (2007) Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. *Science* 315:1860–1862.
2. Buschman TJ, Miller EK (2009) Serial, covert shifts of attention during visual search are reflected by the frontal eye fields and correlated with population oscillations. *Neuron* 63:386–396.
3. Asaad WF, Eskandar EN (2008) A flexible software tool for temporally-precise behavioral control in Matlab. *J Neurosci Methods* 174:245–258.
4. Asaad WF, Eskandar EN (2008) Achieving behavioral control with millisecond resolution in a high-level programming environment. *J Neurosci Methods* 173:235–240.
5. Schwarz GE (1978) Estimating the dimension of a model. *Ann Stat* 6:461–464.
6. Raftery AE (1995) Bayesian model selection in social research. *Sociol Methodol* 25: 111–163.
7. Cowan N (2001) The magical number 4 in short-term memory: A reconsideration of mental storage capacity. *Behav Brain Sci* 24:87–114, discussion 114–185.
8. Xu Y, Chun MM (2006) Dissociable neural mechanisms supporting visual short-term memory for objects. *Nature* 440:91–95.
9. Fuster JM, Alexander GE (1971) Neuron activity related to short-term memory. *Science* 173:652–654.
10. Funahashi S, Bruce CJ, Goldman-Rakic PS (1989) Mnemonic coding of visual space in the monkey's dorsolateral prefrontal cortex. *J Neurophysiol* 61:331–349.
11. Curtis CE, D'Esposito M (2003) Persistent activity in the prefrontal cortex during working memory. *Trends Cogn Sci* 7:415–423.
12. Todd JJ, Marois R (2005) Posterior parietal cortex activity predicts individual differences in visual short-term memory capacity. *Cogn Affect Behav Neurosci* 5: 144–155.
13. Todd JJ, Marois R (2004) Capacity limit of visual short-term memory in human posterior parietal cortex. *Nature* 428:751–754.
14. Palva JM, Monto S, Kulashekhar S, Palva S (2010) Neuronal synchrony reveals working memory networks and predicts individual memory capacity. *Proc Natl Acad Sci USA* 107:7580–7585.
15. Toth LJ, Assad JA (2002) Dynamic coding of behaviourally relevant stimuli in parietal cortex. *Nature* 415:165–168.

16. Bichot NP, Schall JD, Thompson KG (1996) Visual feature selectivity in frontal eye fields induced by experience in mature macaques. *Nature* 381:697–699.
17. Barash S, Bracewell RM, Fogassi L, Gnadt JW, Andersen RA (1991) Saccade-related activity in the lateral intraparietal area. I. Temporal properties; comparison with area 7a. *J Neurophysiol* 66:1095–1108.
18. Bruce CJ, Goldberg ME (1985) Primate frontal eye fields: I. Single neurones discharging before saccades. *J Neurophysiol* 53:607–635.
19. Olejnik S, Algina J (2003) Generalized eta and omega squared statistics: Measures of effect size for some common research designs. *Psychol Methods* 8:434–447.

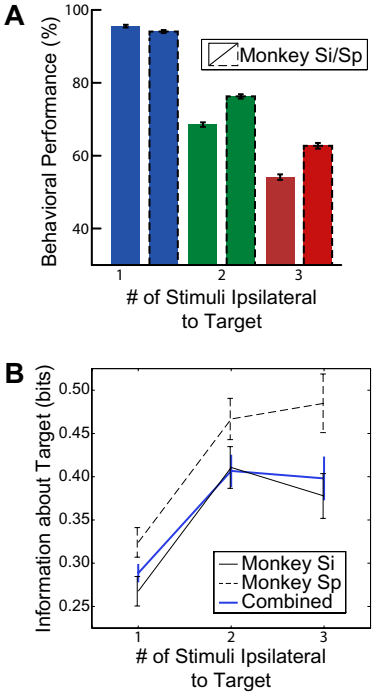

Fig. S1. (*A*) Behavioral performance for both monkeys. (*B*) Information in each hemisphere for both monkeys.

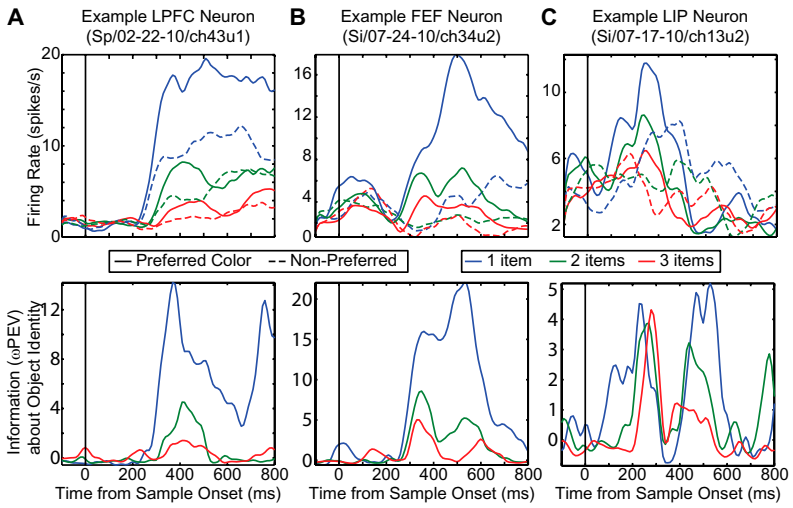

Fig. S2. Example (*A*) LPFC, (*B*) FEF, and (*C*) LIP neurons. (*Upper*) The firing rate response to preferred and nonpreferred colors at each neuron's preferred object location. (*Lower*) The bias-correct percentage of explained variance (ωPEV) derived from firing rate response.

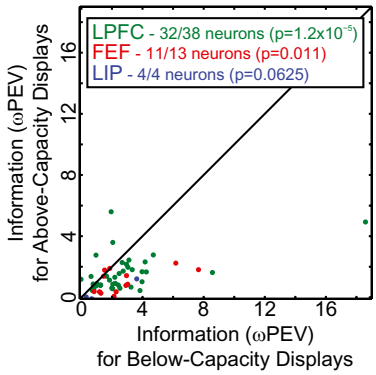

Fig. S3. Distribution of information for stimuli when the entire display is below (*x* axis) or above (two and three items, *y* axis) capacity. Only neurons that are selective for stimulus identity in both conditions are presented. Significance is determined by a binomial test.